Étude Expérimentale d'Extraction d'Information dans des Retranscriptions de Réunions

Pegah Alizadeh, Peggy Cellier, Thierry Charnois, Bruno Crémilleux, Albrecht Zimmermann

Étude Expérimentale d'Extraction d'Information dans des Retranscriptions de Réunions

Pegah ALIZADEH [1] Peggy CELLIER [2] Thierry CHARNOIS [3]
Bruno CRÉMILLEUX [1] Albrecht ZIMMERMANN[1]
(1) Normandie Univ., UNICAEN, ENSICAEN, CNRS – UMR GREYC, Caen, France
(2) Univ Rennes, CNRS, INSA, IRISA - UMR 6074, F-35000 RENNES, FRANCE
(3)LIPN-UMR CNRS 7030, PRES SORBONNE PARIS-CITÉ, FRANCE
peggy.cellier@irisa.fr, thierry.charnois@lipn.univ-paris13.fr,
{pegah.alizadeh, bruno.cremilleux, albrecht.zimmermann}@unicaen.fr

RÉSUMÉ

Nous nous intéressons dans cet article à l'extraction de thèmes à partir de retranscriptions textuelles de réunions. Ce type de corpus est bruité, il manque de formatage, il est peu structuré avec plusieurs locuteurs qui interviennent et l'information y est souvent éparpillée. Nous présentons une étude expérimentale utilisant des méthodes fondées sur la mesure *tf-idf* et l'extraction de *topics* sur un corpus réel de référence (le corpus AMI) pour l'étude de réunions. Nous comparons nos résultats avec les résumés fournis par le corpus.

ABSTRACT

An Experimental Approach For Information Extraction in Multi-Party Dialogue Discourse

In this paper, we address the task of information extraction for meeting transcripts. The meeting documents are not usually well-structured and lacks of formatting and punctuation while the information are distributed over multiple sentences. We investigate on the use of numerical statistic or topic modeling methods on a real dataset containing multi-part dialogue texts. We evaluate our experiments with respect to the summaries provided in the dataset.

MOTS-CLÉS : Extraction d'information, corpus de dialogue, détection de thèmes.

KEYWORDS: Information Extraction, Dialogue Texts, Topic Modeling.

1 Introduction

La plupart des gens passent une grande quantité de leur temps en réunion. Une fois celle-ci terminée, il reste encore un compte rendu à produire, rendant compte des principaux aspects abordés lors de la réunion, comme les problèmes rencontrés et les décisions prises. Il est aujourd'hui possible d'enregistrer et de stocker une réunion en audio voire en vidéo. À partir de ces enregistrements, plusieurs outils dits *"Speech-to-text"* [1] permettent ensuite de générer une retranscription textuelle de tout ce qui a été dit lors de la réunion. Un enjeu important est alors d'être capable d'extraire automatiquement de ces retranscriptions textuelles, souvent fortement bruitées, des informations et des résumés facilitant la création du compte rendu de la réunion. Le projet REUs (relevant de l'appel

1. Exemple d'outil : http ://www.vocapia.com

à projets FUI 22), dans lequel s'inscrit ce travail, a pour but ultime de générer automatiquement un compte-rendu de réunion entre plusieurs humains à partir de la retranscription textuelle de celle-ci.

Quelques travaux se sont intéressés à l'extraction d'information dans les retranscriptions de réunions. Le système le plus abouti est celui présenté par (Tur *et al.*, 2010, 2008) qui fournit un système d'analyse de réunions appelé CALO. CALO retranscrit automatiquement les minutes de la réunion en texte puis différentes parties de la réunion sont identifiées et annotées : les thèmes, les tours de parole, les actions, les décisions et deux résumés (abstractif et extractif) sont fournis. Pour détecter les thèmes et segmenter le texte, CALO s'appuie sur un modèle génératif de thèmes proche de la méthode LDA (Latent Drichlet Allocation) (Purver *et al.*, 2006). À l'aide d'une approche non supervisée, une segmentation de la réunion est produite et permet de connaître simultanément de "quoi" les participants parlent (i.e., les thèmes) à la réunion et "quand" (i.e., segments). Malgré une approche non-supervisée, la méthode nécessite de faire des hypothèses sur la distribution des thèmes et le nombre de segments dans le texte de la réunion. Ce genre d'hypothèses limite l'utilisation de l'approche sur des applications réelles. En ce qui concerne l'extraction d'actions et de décisions, CALO utilise une approche structurelle. Les différentes prises de paroles de la réunion sont classées en fonction de leur rôle dans le processus : "définition de tâche", "accord" et "acceptation d'une responsabilité". Puis les actions (Purver *et al.*, 2007) et les décisions (Fernández *et al.*, 2008) sont détectées. Enfin en ce qui concerne la tâche de génération d'un résumé, CALO extrait différentes versions réduites du texte et choisit celle qui a la meilleure valeur de la mesure ROUGE (Riedhammer *et al.*, 2008) par rapport à des résumés donnés par un oracle.

Dans la littérature, d'autres approches se sont intéressées à ces problématiques ou de façon plus générale à l'analyse de discours. Il existe des approches de segmentation de texte comme (Galley *et al.*, 2003; Georgescul *et al.*, 2007) qui s'appuient sur les changements de distribution lexicale. D'autres approches comme (R. Fernández & Peters, 2008) extraient des mots importants par rapport à un problème en utilisant des classifieurs et des modèles de séquences et en s'appuyant sur les informations lexicales. Les approches de segmentation ou d'extraction de mots sont des méthodes intéressantes mais nécessitent ensuite de déterminer le type de chaque segment (thèmes, décisions, tours de parole, etc). De façon plus générale, plusieurs travaux s'intéressent à la détection d'événements notamment dans les réseaux sociaux (Sayyadi *et al.*, 2009; He *et al.*, 2007). Toutefois les retranscriptions de réunions n'ont pas les mêmes caractéristiques que les textes issus des réseaux sociaux, en particulier dans le rapport au temps. Beaucoup de tweets sont produits par unité de temps dans les réseaux sociaux, alors qu'une seule intervention est produite par tour de parole dans une réunion.

Dans le contexte du projet REUs, afin d'alimenter le compte-rendu devant être produit, plusieurs informations doivent être extraites comme des événements, des décisions, des problèmes, des solutions, un résumé. Les difficultés pour construire un tel système sont multiples. Par exemple, comment détecter qu'une décision importante par rapport à la réunion est prise ? Comment associer une information temporelle à un sujet d'une réunion ou mettre en relation différentes facettes d'un même événement ? Ces tâches ne sont pas simples et elles sont encore plus aiguës dans le contexte de retranscriptions de réunions car les données sont bruitées, les propos ne sont pas toujours structurés, plusieurs locuteurs interviennent rendant plus complexe le fil des propos. Dans cet article, nous nous intéressons à une des étapes nécessaire à la réalisation d'un système générant un compte-rendu de réunions : l'extraction, à partir des retranscriptions textuelles, des thèmes (en anglais *topics*) qui ont été discutés pendant la réunion. Plus particulièrement, nous présentons une étude expérimentale de méthodes classiques d'extraction d'information et de topics que nous avons menée sur le corpus AMI (Carletta, 2007; AMI, 2010), un corpus de référence pour l'étude des réunions. Nous discutons des différentes particularités de ce type de corpus (interactions multiples très variables, flux de

mots non structurés, ponctuation faible, formatage, etc) et de leurs conséquences sur le processus d'extraction d'information.

2 Extraction d'information dans une retranscription de réunion

L'objectif est d'extraire des informations utiles pour un compte-rendu de réunion. Nous nous sommes donc intéressés aux approches permettant d'extraire des mots "importants" dans un texte afin de voir leur comportement sur des retranscriptions de réunions. Nous avons considéré deux approches bien connues : la mesure *tf-idf* et une approche de *topic modeling*, qui sont rappelées ici.

Fréquence du terme - Fréquence inverse du document (*tf-idf*) La Fréquence du terme - Fréquence inverse du document (*tf-idf*) est une technique qui donne les poids les plus élevés aux termes (i.e. mots) qui apparaissent les plus fréquemment dans un document par rapport aux autres documents du corpus. La fréquence d'un terme (tf(w, d)) est simplement le nombre d'occurrences du mot w dans le document d. La fréquence inverse du document (idf) est la proportion de documents du corpus dans laquelle le mot w apparaît. Plus précisément : $idf(w, D) = \log \frac{N}{1+|d \in D \ s.t. \ w \in d|}$ où N est le nombre total de documents du corpus. Le dénominateur est le nombre de documents dans lesquels un mot w apparaît dans le corpus D. La mesure *tf-idf* est ainsi définie par : tfidf$(w, d, D) =$ tf(w, d)idf(w, d)

***Topic Modeling* : Factorisation par matrices non négatives** Il existe plusieurs approches de *topic modeling* permettant de définir des thèmes apparaissant dans un corpus. Il y a des approches probabilistes telles que l'approche LDA (Latent Dirichlet Allocation) (Blei *et al.*, 2003) mais aussi des méthodes s'appuyant sur l'agèbre linéaire comme *NMF* (Factorisation par matrices non négatives) (Lee & Seung, 1999). Dans cet article nous avons choisi d'utiliser la méthode *NMF* car d'autres expériences non décrites ici ont montré que *NMF* donnait dans ce cas de meilleurs résultats. Dans l'approche *NMF*, on représente m mots apparaissant dans n documents via une matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$. L'objectif de la méthode *NMF* est ensuite de décomposer $\mathbf{A}$ en deux matrices $\mathbf{W}$ et $\mathbf{H}$ telles que $\mathbf{A} \sim \mathbf{WH}$. Les colonnes de la matrice $\mathbf{W} \in \mathbb{R}^{m \times k}$ sont les thèmes (*topics*) identifiés et les lignes les mots représentant ces thèmes. Chaque case de la matrice représente le poids pour un mot d'être relié à un thème. La matrice $\mathbf{H} \in \mathbb{R}^{k \times n}$ représente la façon dont un document est en relation avec k thèmes.

3 Corpus AMI

Le corpus de réunions AMI (Carletta, 2007) contient 100 heures de réunion enregistrées en utilisant plusieurs appareils d'enregistrement synchronisés. Tous les participants aux réunions parlent anglais. Pour certains l'anglais est leur langue maternelle et pour d'autres non. Cela représente 171 réunions qui se regroupent en deux types : les **réunions à base de scénario** (*scenario-based meetings*) et les **réunions sans scénario** (*non-scenario based meetings*). Les **réunions à base de scénario** font partie de séries de 3 à 4 réunions simulant les réunions de quatre participants devant concevoir un système. Les réunions d'une même série ont toutes lieu le même jour. Les **réunions sans scénario** sont de vraies réunions qui ont été enregistrées. Le corpus AMI a été transcrit et annoté avec des informations concernant les entités nommées présentes et les différents tours de parole. Pour chaque réunion un résumé abstractif et un résumé extractif sont fournis. Le résumé abstractif d'une réunion contient environ 200 mots et consiste en un texte libre donnant un résumé général de la réunion ainsi que des

explications à propos des décisions prises et des problèmes évoqués lors de la réunion. Le résumé extractif identifie des morceaux du texte de la réunion qui couvrent les informations contenues dans le résumé abstractif. Notons qu'un élément du résumé abstractif peut faire référence à plusieurs tours de parole et qu'un tour de parole peut être relié à plusieurs éléments du résumé abstractif.

Disponibilité du corpus Une version segmentée du corpus AMI (Carletta, 2007; AMI, 2010) est disponible en ligne [2]. Toutefois le corpus mis à disposition est au format NXT (Nite XML Toolkit), qui ne permet pas de faire des traitements sur le texte. Nous avons recréé les fichiers textes de retranscription et les avons mis à disposition en ligne [3].

4 Extraction d'information dans le corpus AMI

Procédure d'expérimentation Comme mentionné précédemment, le corpus AMI contient deux types de meetings : les réunions avec scénario et sans scénario. Nous nous sommes concentrés sur les réunions avec scénario car elles représentent 138 réunions sur 171. De plus, la plupart des réunions sans scénario n'ont pas de résumé disponible, ce qui ne permet pas d'évaluation.

Pour chacune des deux expériences nous appliquons une approche : *tf-idf* ou *NMF*. Le résultat de l'approche est ensuite évalué en comparant les informations extraites des réunions aux informations présentes dans les deux types de résumé (abstractif et extractif). La précision par rapport à un résumé abstractif (respectivement extractif) est donnée par la formule suivante : $\frac{\|\{w \in R\} \cap \{w \in S\}\|}{\|\{w \in R\}\|}$ où R est l'ensemble des mots du résultat et S les mots du résumé considéré (abstractif ou extractif). Notons que l'on applique une racinisation sur les mots et c'est la racine qui est prise en compte pour l'appariement de mots. Par exemple, "painting" et "paint" sont considérés comme le même mot.

4.1 Comparaison entre séries de réunions

Dans cette première expérience, nous regroupons toutes les réunions d'une même série en un seul corpus. Il y a donc 34 corpus, chacun représentant une série de réunions. Ensuite nous extrayons les informations propres à chacun des corpus selon les deux méthodes.

tf-idf Avec l'approche utilisant la mesure *tf-idf*, nous calculons les 20 mots les plus importants pour chaque série de réunions. La table 1 (première colonne) donne des exemples de mots extraits avec cette mesure pour 2 séries de réunions avec scénario [4]. On voit que pour la série 1 le terme avec le plus haut score par rapport aux autres mots de la réunion est "matthew", suivi de "mael".

À la figure 1, le schéma du haut donne le score de précision par rapport aux résumés par abstraction et par extraction. Comme attendu, la précision pour le résumé extractif (plus de 60%) est meilleure que pour le résume abstractif (moins de 40%). En effet, cette mesure calcule le pourcentage de mots présents dans chaque résumé. Le résumé extractif est plus long et est un sous-ensemble du texte initial contrairement au résumé abstractif qui est une reformulation de ce qui a été dit dans la réunion.

2. http://groups.inf.ed.ac.uk/ami/download/
3. https://github.com/pegahani/AMI-prep
4. Tous les résultats pour toutes les séries sont disponibles ici : https://github.com/pegahani/Event_detection/blob/master/result/result_4_block_scen.txt

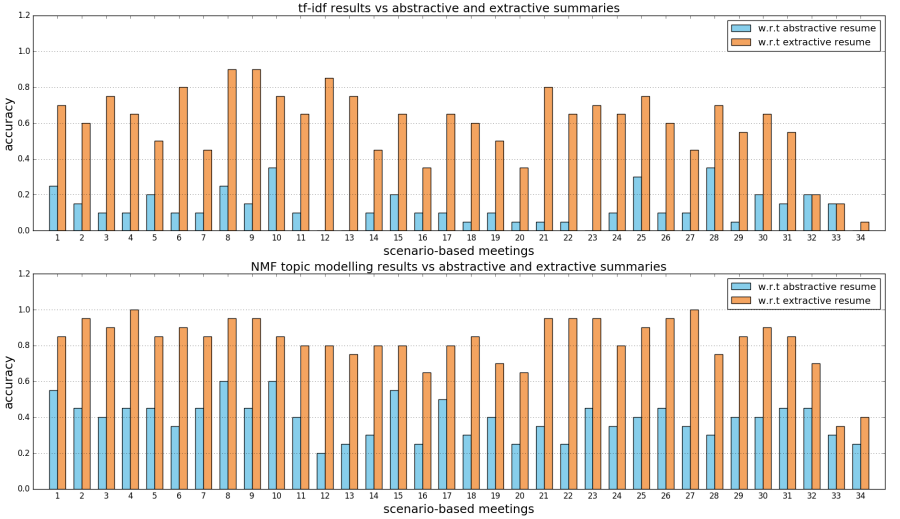

FIGURE 1 – Le graphique du haut montre la précision de l'approche *tf-idf* par rapport aux résumés abstractif et extractif. Le graphique du bas montre la précision de l'approche *NMF* par rapport aux résumés abstractif et extractif.

Série de réunions	Mots extraits avec *tf-idf*	Mots extraits avec *NMF*
Série 1	**'matthew', 'mael', 'anna'**, 'exce', 'ip', 'doctor', 'decline', 'assemble', 'streamed', 'customizing', 'asian', 'voter', 'undes', 'nanne', 'highperformance', 'fik', 'protec', 'underlie', 'provin', 'zebras'	keys, **matthew**, browse, innovation, functionalities, v_c_r_, **mael, anna**, sixteen, perfect, demographic, r_c_, receiver, surf, present blinking, cents, store, movie, presented
Série 4	**'mushroom', 'jordan'**, 'coarse', 'baba', 'alimentation', 'mush', 'gestures', 'kemy', 'institute', 'frahan', 'florent', 'laser', 'longmund', 'sleeping', 'ada', 'saucer', 'trois', 'ecological', 'hmmm', 'eatable'	controller, **mushroom**, gesture, google, pineapple, powerful, david, base, wireless, traditional, lemon, **jordan**, wooden, sophisticated, wire, vocal, participant, ball, bulb, recognise

TABLE 1 – Informations extraites pour 2 séries de réunions avec scénario.

Topic Modeling Nous avons aussi mené une expérience avec une approche *topic Modeling*. Pour cela, nous avons utilisé la méthode *NMF* (cf section 2) pour affecter un thème à chaque réunion avec scénario. Comme nous l'avons signalé précédemment dans cette section, il y a 34 réunions de ce type, nous utilisons donc le modèle *NMF* en fixant à 34 le nombre de thèmes à classifier. Pour implémenter l'approche *NMF*, nous utilisons le package gensim (Řehůřek & Sojka, 2010). Après l'application de *NMF* sur les réunions, un thème est affecté à chaque réunion. Ce thème est représenté par les 20 mots les plus "importants". Sur le tableau 1, les mots extraits pour 2 réunions sont indiqués dans la colonne de gauche.

Comme pour la méthode s'appuyant sur le *tf-idf*, dans le graphe du bas de la figure 1 la précision de l'approche *NMF* sur les 34 corpus est donnée. Nous constatons que l'approche *NMF* donne de meilleurs résultats que la comparaison soit faite avec le résumé abstractif ou extractif. Nous pensons que les meilleurs résultats de *NMF* viennent du fait que dans cette approche tous les mots de toutes les réunions sont considérés pour l'extraction des thèmes. Avec l'approche s'appuyant sur *tf-idf* seuls les mots de la série de réunions qui sont peu fréquents dans les autres réunions vont être extraits. On peut ainsi manquer des mots/thèmes importants.

M_1	M_2	M_3	M_4
25 : animal	9 : age, lunch, teletext	8 : solar, wood	17 : criteria
19 : cat	8 : percent, young	6 : titanium	13 : seven
12 : favourite	6 : pay	5 : spongy, concepts	12 : evaluation
11 : tool, dog	5 : settings, zap, seventy, users	4 : dark, sample, doublecurved, materials, sensor, banana, circuit, cases, vegetables, fruit	9 : sample
7 : training, draw	4 : set, messages, group, mode, infrared, recognition, speech		8 : special, false
6 : rabbit, profit, fish			6 : evaluate, process
4 : friendly, bird, tail, whiteboard, width, characteristics, elephant, morning			5 : prototype, leadership, creativity, under
			4 : scale, single, fifteen, team, average, budget, curve

TABLE 2 – Mots les plus souvent répétés et ayant les plus hauts scores *tf-idf*.

4.2 Caractérisation de réunions individuelles

Comme indiqué précédemment, les réunions d'AMI sont destinées à simuler un projet, par exemple la conception d'une télécommande, du début à la fin. Toutes les réunions liées à une série (c'est-à-dire des réunions jouées par le même groupe de participants) sont condensées en une seule journée. Il y a 34 séries différentes. Les différentes séries de réunions portent sur le même scénario mais varient dans la mise en œuvre concrète du script. Si les réunions d'une même série comportent des caractéristiques communes, nous nous attendons à ce qu'elles apparaissent dans les termes extraits par *tf-idf*.

Pour vérifier notre hypothèse : nous traitons une série de réunions comme un corpus et identifions pour chaque réunion individuelle dans cette série les 20 mots avec les scores plus élevés selon *tf-idf*. Pour chacun de ces mots, nous comptons alors le nombre de fois où il est apparu dans l'ensemble des mots dérivés de la même étape dans les autres séries. Par exemple, nous rassemblons les ensembles de mots dérivés de la première réunion de chaque série (M_1) et comptons les doublons. La table 2 montre les mots répétés plus que 4 fois pour chaque étape (10% arrondi)[5]. On voit que le terme «animal» a été utilisé dans 73.5% de toutes les premières réunions.

La table 2 (y compris 10% des vingt mots les plus élevés pour l'ensemble des 34 séries) montre que la plupart des mots sont apparus dans la première ou la dernière réunion de la série alors que les deux réunions intermédiaires ont une petite part. Par exemple, dans l'étape M_1, «animal» est répété 29 fois, mais dans l'étape M_2, moins de mots apparaissent, par exemple «age» est apparu 9 fois. Nous voyons que la première et la dernière réunions d'une série ont des mots plus liés que les deuxième et troisième réunions. Cela montre que la première rencontre semble toujours être un brainstorming lié aux animaux et la quatrième réunion sur la façon d'évaluer le succès du projet. Lors de la troisième réunion, la discussion a porté sur les matériaux à utiliser. Enfin, la présence de «lunch» en tant que mot le plus répété lors de la deuxième réunion montre que cette réunion a eu lieu en fin de matinée.

4.3 Comparaison entre réunions d'une même série

Dans une seconde expérience, nous extrayons les informations propres à une réunion par rapport aux autres réunions de la même série. Nous avons testé les deux méthodes : *tf-idf* et *NMF*. Contrairement à l'expérience précédente, la méthode s'appuyant sur *tf-idf* a donné dans nos expériences de meilleurs résultats que celle s'appuyant sur *NMF*. Nous supposons que cela ait du au fait que les corpus de cette expérience sont plus courts.

5. Pour les résultats complets, voir : `https://github.com/pegahani/Event_detection/blob/master/result/result_4_4.txt`

Aux figures 2 et 3, les scores de précision obtenus pour chacune des réunions de chacune des séries par rapport au résumé abstractif ou extractif sont donnés. Si l'on compare les résultats par rapport aux deux types de résumés, on retrouve la même tendance que sur l'expérience précédente portant sur les séries de réunion (section 4.1) : les scores des résumés par extraction sont meilleurs (la plupart au dessus de 0.6) que pour les résumés par abstraction (au dessous de 0.6).

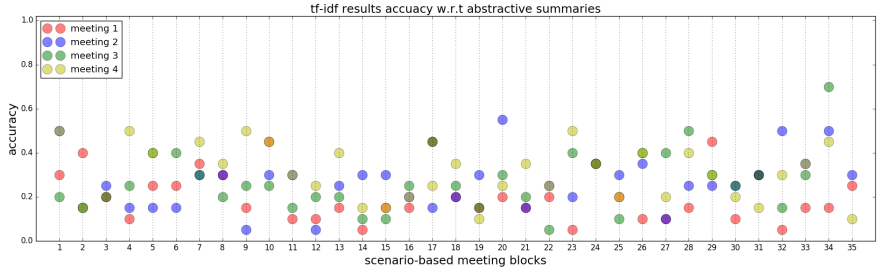

FIGURE 2 – Précision, par rapport au résumé abstractif, des thèmes extraits par *tf-idf*.

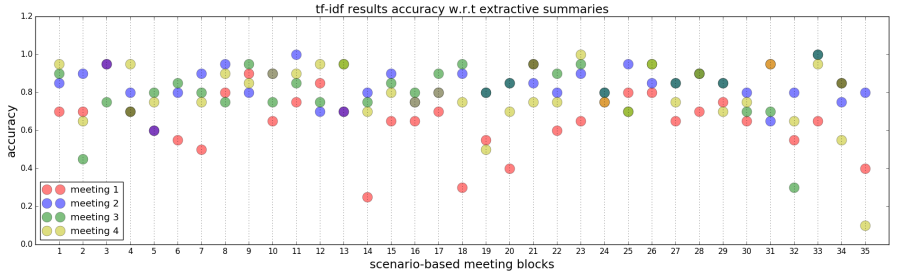

FIGURE 3 – Précision, par rapport au résumé extractif, des thèmes extraits par *tf-idf*.

5 Conclusion

Ce travail représente une étude de base. L'objectif est d'alimenter un outil de génération automatique de compte-rendu de réunion. Nous avons mené des expériences sur l'extraction d'information dans des corpus de transcription de réunions en testant deux approches de l'état-de-l'art sur un corpus de référence AMI. Les résultats obtenus nous semblent encourageants sur ce type de corpus et justifient l'utilisation d'approches robustes. Ils montrent aussi que les deux approches sont complémentaires puisque l'approche *topic modeling* (*NMF*) donne de meilleurs résultats pour la caractérisation de réunions individuelles, alors que pour la comparaison de réunions d'une même série c'est l'approche *tf-idf*. Nous souhaitons par la suite travailler sur des méthodes permettant d'affiner l'information extraite.

Remerciements Ce travail est soutenu par le FUI 22 (projet REUs).

Références

AMI (2010). Augmented multi-party interaction. `http://www.amiproject.org`. [Online].

BLEI D. M., NG A. Y. & JORDAN M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, **3**, 993–1022.

CARLETTA J. (2007). Unleashing the killer corpus : experiences in creating the multi-everything AMI meeting corpus. *Language Resources and Evaluation*, **41**, 181–190.

FERNÁNDEZ R., FRAMPTON M., EHLEN P., PURVER M. & PETERS S. (2008). Modelling and detecting decisions in multi-party dialogue. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, SIGdial '08, p. 156–163, Stroudsburg, PA, USA : Association for Computational Linguistics.

GALLEY M., MCKEOWN K., FOSLER-LUSSIER E. & JING H. (2003). Discourse segmentation of multi-party conversation. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1*, ACL '03, Stroudsburg, PA, USA : Association for Computational Linguistics.

GEORGESCUL M., CLARK A. & ARMSTRONG S. (2007). Exploiting structural meeting-specific features for topic segmentation. In *TALN/RECITAL*, p. 15–24, Toulouse (France).

HE Q., CHANG K. & LIM E.-P. (2007). Analyzing feature trajectories for event detection. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, p. 207–214, New York, NY, USA : ACM.

LEE D. D. & SEUNG H. S. (1999). Learning the parts of objects by nonnegative matrix factorization. *Nature*, **401**, 788–791.

PURVER M., DOWDING J., NIEKRASZ J., EHLEN P., NOORBALOOCHI S. & PETERS S. (2007). Detecting and summarizing action items in multi-party dialogue. In *In Proc. of the 9th SIGdial Workshop on Discourse and Dialogue*.

PURVER M., GRIFFITHS T. L., KÖRDING K. P. & TENENBAUM J. B. (2006). Unsupervised topic modelling for multi-party spoken discourse. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics*, ACL-44, p. 17–24, Stroudsburg, PA, USA : Association for Computational Linguistics.

R. FERNÁNDEZ, M. FRAMPTON J. D. A. A. P. E. & PETERS S. (2008). Identifying relevant phrases to summarize decisions in spoken meetings. In *Proceedings of Interspeech'08*, Brisbane.

ŘEHŮŘEK R. & SOJKA P. (2010). Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, p. 45–50, Valletta, Malta : ELRA. `http://is.muni.cz/publication/884893/en`.

RIEDHAMMER K., FAVRE B. & HAKKANI-TÜR D. (2008). Packing the Meeting Summarization Knapsack. In *Interspeech, Brisbane (Australia)*, Unknown, Unknown or Invalid Region.

SAYYADI H., HURST M. & MAYKOV A. (2009). Event detection and tracking in social streams. In *In Proceedings of the International Conference on Weblogs and Social Media (ICWSM 2009). AAAI*.

TUR G., STOLCKE A., VOSS L., DOWDING J., FAVRE B., FERNANDEZ R., FRAMPTON M., FRANDSEN M., FREDERICKSON C., GRACIARENA M., HAKKANI-TUR D., KINTZING D., LEVEQUE K., MASON S., NIEKRASZ J., PETERS S., PURVER M., RIEDHAMMER K., SHRIBERG E., TIEN J., VERGYRI D. & YANG F. (2008). The calo meeting speech recognition and understanding system. In *2008 IEEE Spoken Language Technology Workshop*, p. 69–72.

TUR G., STOLCKE A., VOSS L., PETERS S., HAKKANI-TUR D., DOWDING J., FAVRE B., FERNANDEZ R., FRAMPTON M., FRANDSEN M., FREDERICKSON C., GRACIARENA M., KINTZING D., LEVEQUE K., MASON S., NIEKRASZ J., PURVER M., RIEDHAMMER K., SHRIBERG E., TIEN J., VERGYRI D. & YANG F. (2010). The calo meeting assistant system. *IEEE Transactions on Audio, Speech, and Language Processing*, **18**, 1601–1611.